3D Pareto Front: Cost-Recall-Latency Trade-off